Source Address Validation in Intra-domain Networks

draft-cheng-savnet-intra-domain-sav-igp-03
draft-cheng-savnet-intra-domain-sav-bgp-01

Presenter : Shengnan Yue (China Mobile)

Co-authors: Weiqiang Cheng (China Mobile)
Dan Li (Tsinghua University)
Changwang Lin (New H3C Technologies)
Shengnan Yue (China Mobile)

IETF121

Contents

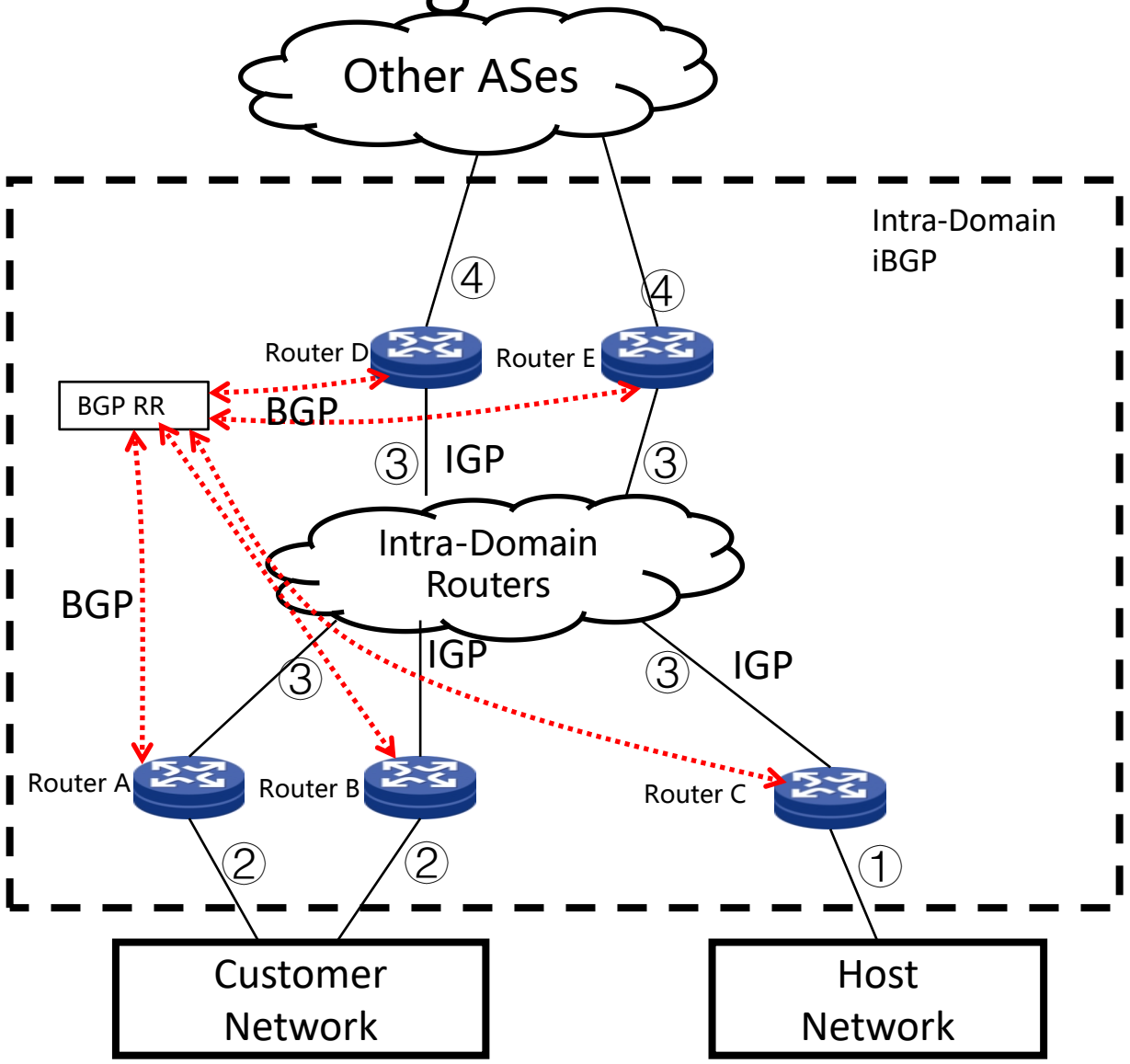
- SAV on Access, Intermediate, Ext Interface

  Remove the SAV source prefix Flag.

- Intra-Domain SAV via BGP
- Running Code

  Conduct experiments and test data in the LAB.

Networking

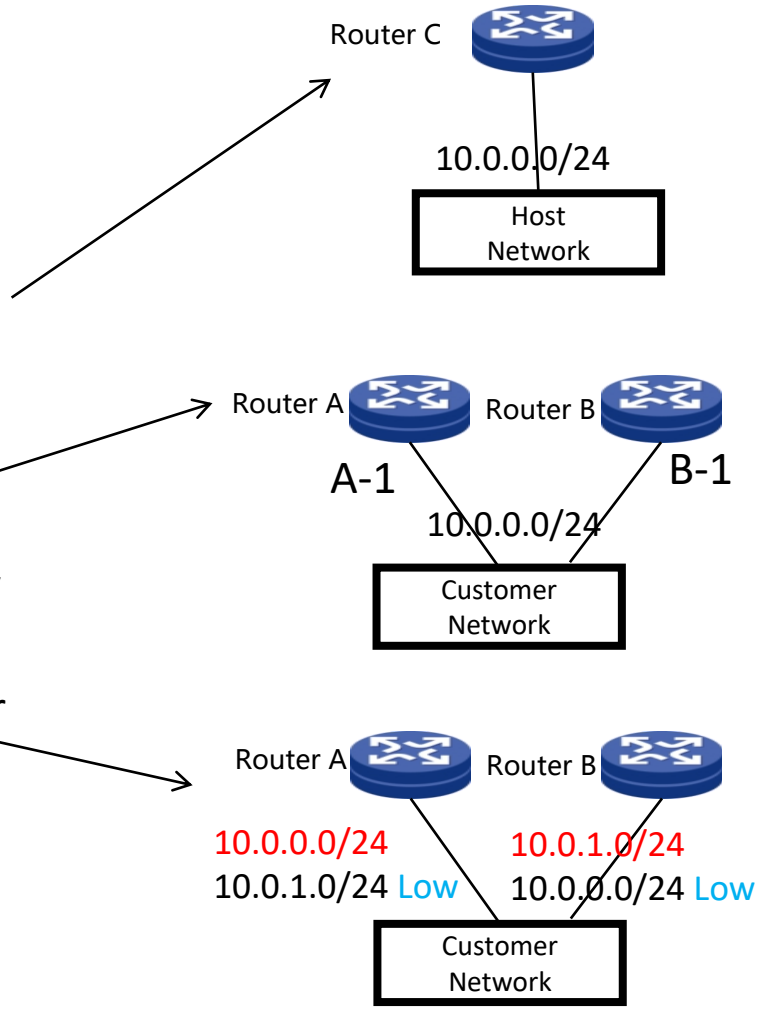


- Three types of device roles:
  - Access devices, such as Router A, Router B,Router C;
  - Intermediate devices;
  - Border devices, such as Router D, Router E
- Four types of interfaces:
  ①: Access interface on host-network;
  ②: Access interface on customer-network;
  ③: Intermediate interface within the domain;
  ④: Ext interfaces on AS border.
- Two types of SAV rules:
  - Intra-domain interfaces ①- ③ primarily deploy allow-list rules
  - Ext Interface ④ primarily deploy block-list rules.
- Two Scenarios:
  - Scenario 1: The source prefix is advertised by the IGP protocol;
  - Scenario 2: The source prefix is communicated by the BGP protocol.

  while the IGP protocol is used to calculate the domain's connectivity

SAV on Access Interface

- It is recommended to use Allow-List Rule .
- For the Host Network, add SAV Allow-List Rule directly according to the subnet address.
- For the Customer Network, if the prefiex announcing to Router A and Router B are identical, Routers directly generate SAV rules based on the subnets.
- In multi-homed subnet scenario, generate SAV rules based on high-priority and low-priority prefixes. Different subnets advertise prefixes or configuring the access routers routes.
  High priority 10.0.0.0/24 and Low priority 10.0.1.0/24 are advertise to Router A.
  High priority 10.0.1.0/24 and Low priority 10.0.0.0/24 are advertise to Router B.



SAV Rule on C:

	10.0.0.0/24
Allow-List	C-1

SAV Rule of A & B:

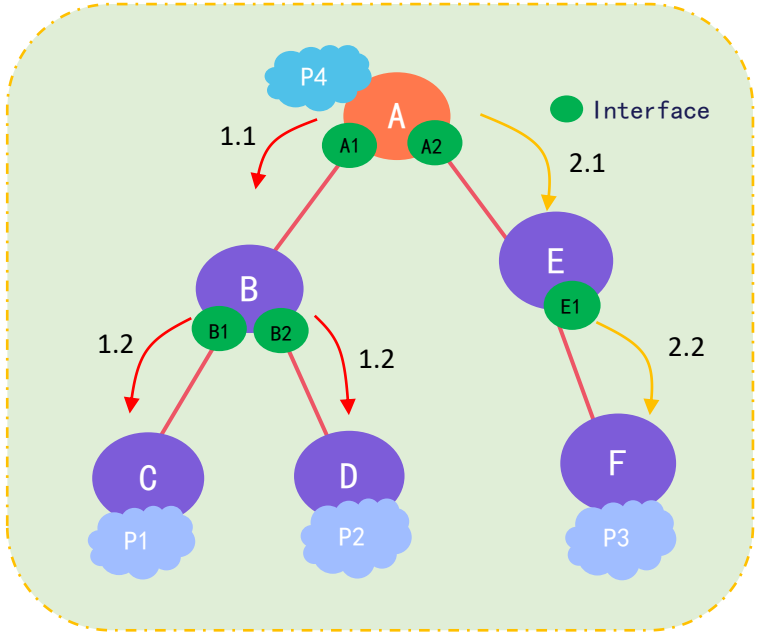
	10.0.0.0/24
Allow-List	A-1
Allow-List	B-1

SAV Rule of A & B:

	10.0.0.0/24	10.0.1.0/24
Allow-List	A-1	A-1
Allow-List	B-1	B-1

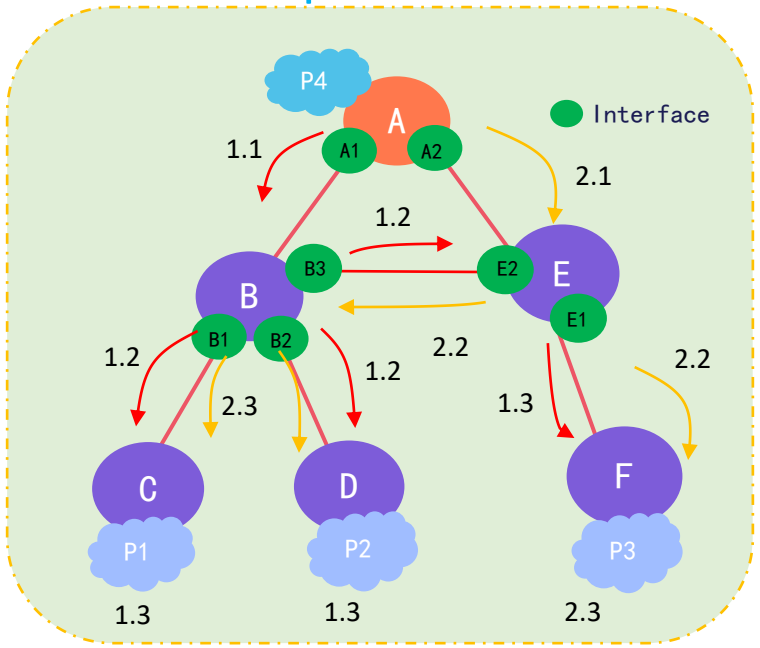
SAV on Intermediate Interface

- It is recommended to use Allow-List Rule.
- Use the LSDB information of the IGP protocol to calculate connectivity and determine the interface.
- Verification on the intermediate interface is optional.



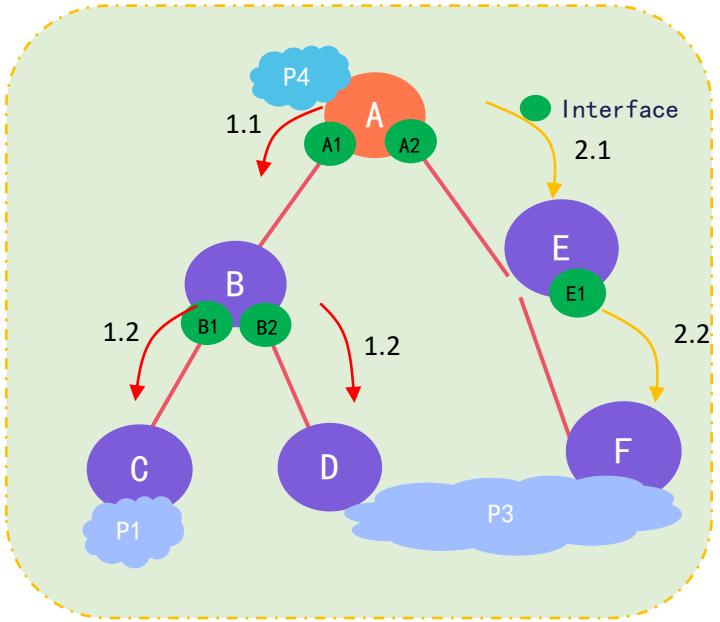
Scenario 1,
SAV Rule of A:

	P1	P2	P3
Allow-List	A1	A1	A2



Scenario 2,
SAV Rule of A:

	P1	P2	P3
Allow-List	A1,A2	A1,A2	A1,A2

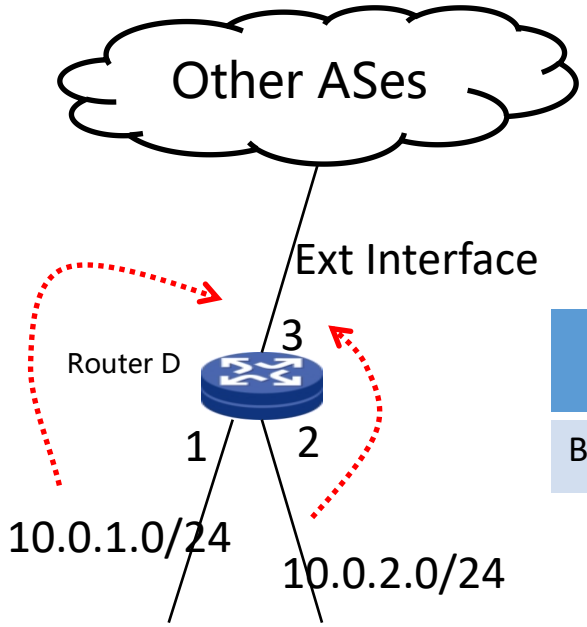


Scenario 3, SAV Rule of A:

	P1	P3
Allow-List	A1	A1,A2

SAV on Ext Interface

- It is recommended to use Block-List Rule for filtering on Ext interface.
- On the border router, first calculate the prefix and Ingress interface using the previously mentioned method for intermediate interfaces, then convert it accordingly into the Block-List of the Ext interface.



	10.0.1.0/24	10.0.2.0/24
Block-List	D-3	D-3

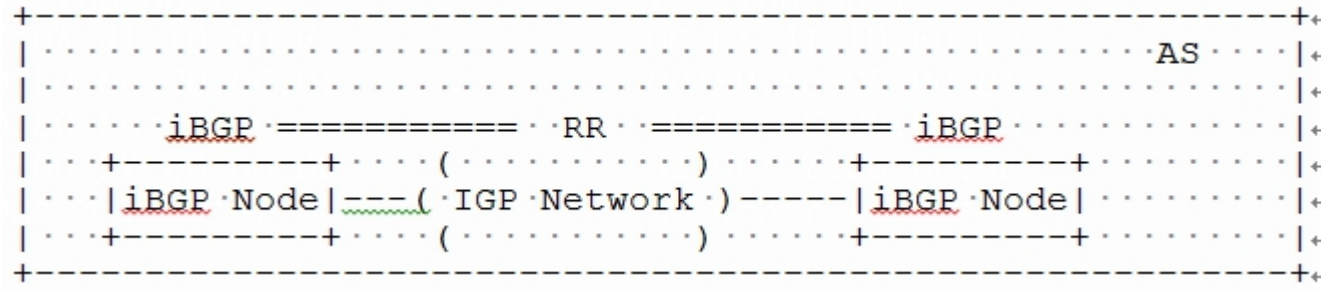
BGP for Intra-Domain SAV

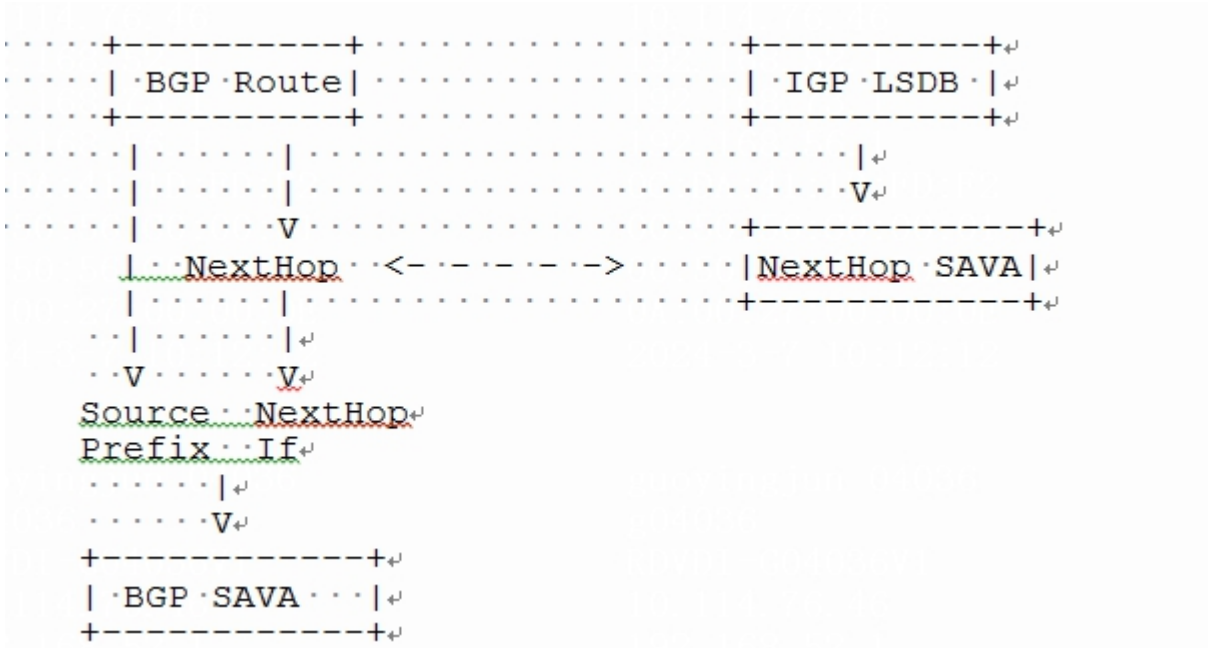
- It is common to propagate source prefixes via iBGP, which establishes neighbors using the underlying IGP network.
- The principle of BGP calculation:

  Step 1: Calculate the SAV rules related to the BGP route gateway using the previously described process of computing SAV rules via IGP.

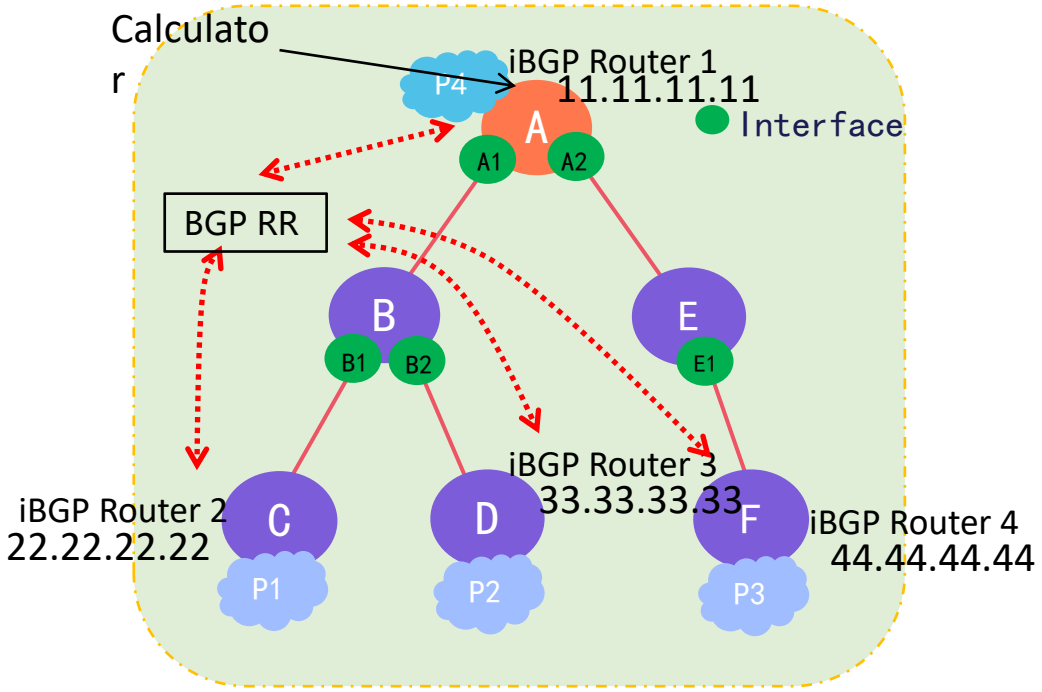
  Step 2: Obtain the source prefixes published by the BGP neighbor and fetch the SAV rule table corresponding to the next hop of the BGP neighbor. Then, inherit the source interface list from the neighbor's SAV rule table to generate the BGP SAV rule table.

- Specific BGP Community can be used to identIfy wheather the source prefixes can (or cannot) be added to SAV entries to prevent the false filtering.





Scenario: BGP Advertise Source Prefix



IGP SAV Rule:

Prefix	22.22.22.22	33.33.33.33	44.44.44.44
Ingress	A1	A1	A2

BGP Prefix:

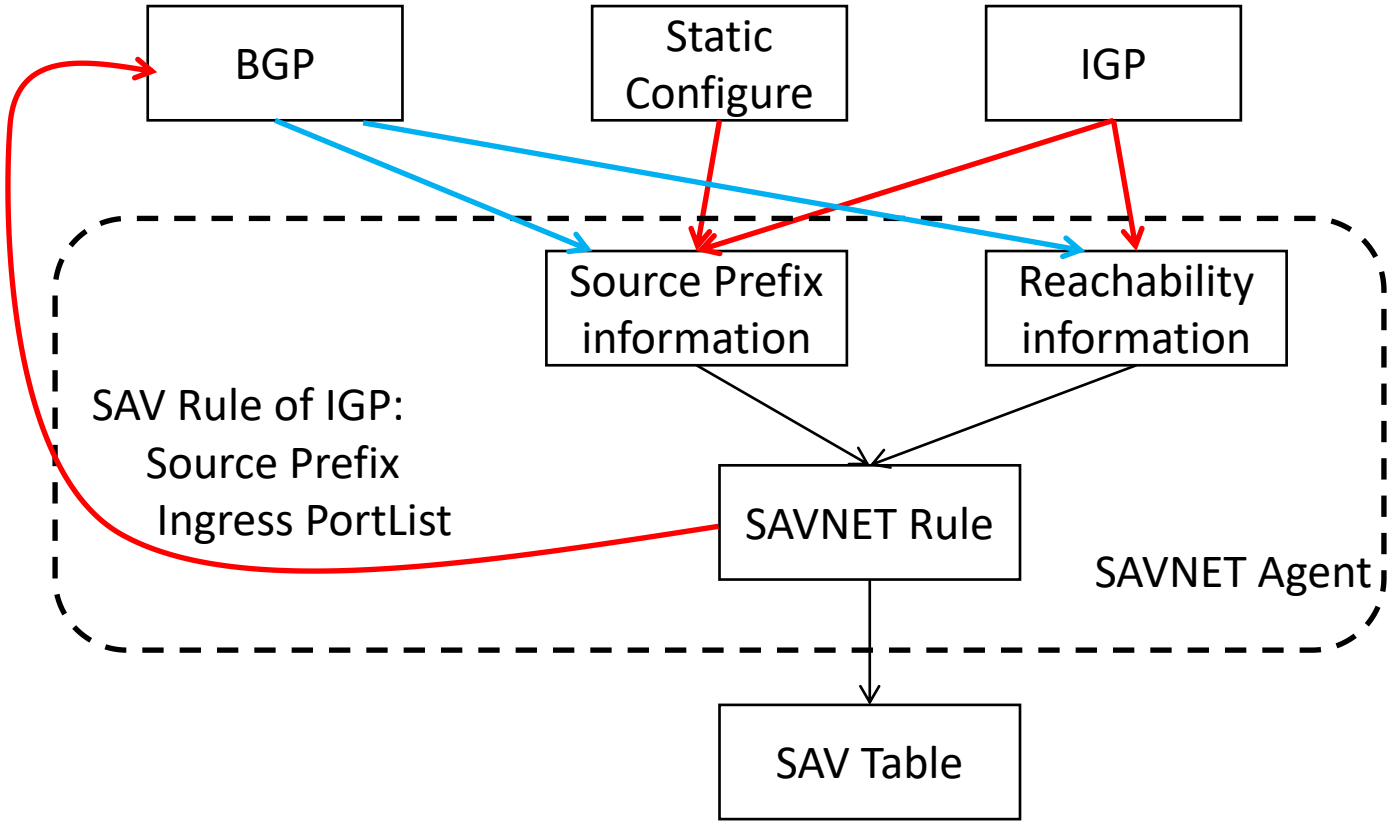
Prefix	P1	P2	P3
Nexthop	22.22.22.22	33.33.33.33	44.44.44.44

BGP SAV Rule:

Prefix	P1	P2	P3
Allow-List	A1	A1	A2

The overall Intra-domain SAV framework

- The source prefix of Intra-domain can be statically configured, or distributed within the domain via the IGP.
- Connectivity calculation is based on the existing IGP LSDB information. During connectivity calculation, it is required to perform two-way neighbor check for validation.
- The source prefix of Intra-domain advertised via iBGP, iterated to the IGP source prefix based on the gateway address, inheriting the IGP SAV Rule's Ingress Portlist.
- The SAVNET Agent utilizes connectivity algorithms to calculate the source port information of the source prefix based on the source prefix information and reachable information, and generates the SAV rule.

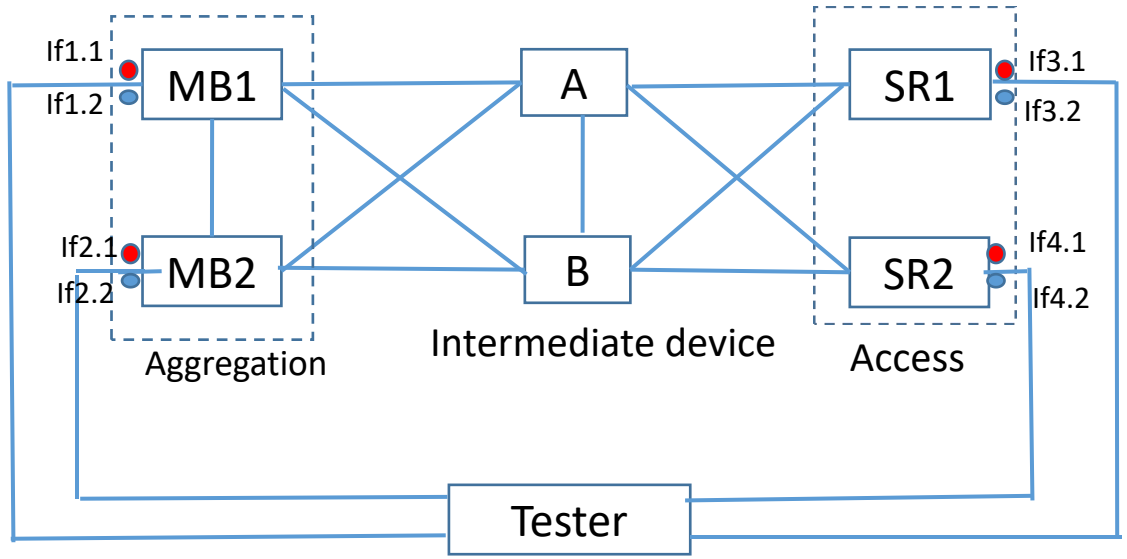


Running Code

- Lab Interop-test Status
  Hardware devices and software implementations which have passed the interoperability tests of intra-domain SAVNET solution in these two drafts hosted by China Mobile in 2024:
  H3C CR16010H-FA and CR16000-M8
  Ruijie RG-N8010-R

Both single-homed and multi-homed subnet scenarios have passed interoperability testing.

SAVNET Test Topology



Results for Performance(1)

- Little impact of Access-side Allow-list on Forwarding Latency

Manufacturer	Route Number	200000	400000
H3C	Forwarding Latency(us)	50.24us	50.09us
	Enable Allow-list Forwarding Latency (us)	50.92us	50.43us
Ruijie	Forwarding Latency(us)	78.71 us	78.72 us
	Enable Allow-list Forwarding Latency (us)	78.73 us	78.72 us

- Little impact of Intermediate Device Down Link Allow-list on Forwarding Latency

Manufacturer	Route Number	200000	400000
H3c	Forwarding Latency(us)	50.24us	50.09us
	Enable Allow-list Forwarding Latency (us)	50.68us	50.52us
Ruijie	Forwarding Latency(us)	78.71 us	78.72 us
	Enable Allow-list Forwarding Latency (us)	78.73 us	78.73 us

Results for Performance(2)

- Little impact of Aggregation-side Up-link Block-list on Forwarding Latency

Manufacturer	Route Number	200000	400000
H3C	Forwarding Latency(us)	51.1us	51.26us
	Enable Block-list Forwarding Latency (us)	52.28us	52.18us
Ruijie	Forwarding Latency(us)	72.93 us	72.94 us
	Enable Block-list Forwarding Latency (us)	72.94 us	72.93 us

- Little impact of Aggregation-side Down-Link Allow-list on convergence

Manufacturer	Route Number	200000	400000
H3C	Normal convergence (ms)	3.53ms	3.29ms
	Allow-list, Direct Route(ms)	4.15ms	4.38ms
	Allow-list, Same Destination Number (ms)	4.55ms	4.21ms
Ruijie	Normal convergence(ms)	3.5ms	4.2ms
	Allow-list, Direct Route(ms)	1.9ms	3.1ms
	Allow-list, Same Destination Number (ms)	2.7ms	4.0ms

Results for Performance(3)

CPU Usage

Manufacturer	Route Number	100000	200000
H3C-MB	Normal	2%	2%
	SAVNET	2%	2%
H3C-SR	Normal	15%	15%
	SAVNET	15%	15%
RJ-MB	Normal	5.6%	5.6%
	SAVNET	5.7%	5.7%
RJ-SR	Normal	5.8%	5.8%
	SAVNET	5.7%	5.6%

Memory

Manufacturer	Route Number	100000	200000
MB-H3C	Normal	36.5%	37.9%
	SAVNET	40.4%	46.6%
SR-H3C	Normal	40.6%	42%
	SAVNET	41.3%	43.4%
RJ-MB	Normal	41.1%	41.3%
	SAVNET	41.5%	42.6%
RJ-SR	Normal	40.6%	41.0%
	SAVNET	40.5%	41.4%

Summary

- Verification also can be optional at the intermediate device, according to the architecture draft.
- Although verification at the intermediate layer may be lenient, in the context of hierarchical and complex networking environments faced by carriers, the connectivity calculation approach does not incorrectly discard legitimate traffic. It is cost-effective, easy to deploy, and has became the suitable technical solution for carriers.

Next Step

- Seeking comments
- Ask for adoption call?

THANKS